cemmap
centre for microdata methods and practice

Factor rotation with non-negativity constraints

Stephen Pudney

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP02/11

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Factor rotation with non-negativity constraints

Stephen Pudney

Institute for Social and Economic Research
University of Essex

This version: January 2, 2011

Abstract

Factor rotation is widely used to interpret the estimated factor loadings from latent variable models. Rotation methods embody a priori concepts of 'complexity' of factor structures, which they seek to minimise. Surprisingly, it is rare for researchers to exploit one of the most common and powerful sources of a priori information: non-negativity of factor loadings. This paper develops a method of incorporating sign restrictions in factor rotation, exploiting a recently-developed test for multiple inequality constraints. An application to the measurement of disability demonstrates the feasibility of the method and the power of non-negativity restrictions.

Keywords: Factor rotation; Inequality constraints; Chen-Szroeter test; Disability

JEL codes: C12, C38, I10

Contact: Steve Pudney, ISER, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; tel. +44(0)1206-873789; email spudney@essex.ac.uk

I am grateful to Marcello Morciano for his help with computation. This work was supported by the Nuffield Foundation (via the project 'The role and effectiveness of disability benefits for older people') and the European Research Council (project no. 269874 [DEVHEALTH]). Support from the Economic and Social Research Council through the MiSoC research centre (award no. RES-518-285-001) and the University of Melbourne through a Faculty Visiting Scholarship in the Department of Economics and Melbourne Institute is also gratefully acknowledged. Data from the English Longitudinal Study of Ageing (ELSA) were developed by researchers based at University College London, the Institute for Fiscal Studies and the National Centre for Social Research and are made available through the UK Data Archive. Neither the collectors of the data nor the UKDA bear any responsibility for the analyses or interpretations presented here.

1 Introduction

Statistical models involving unobservable latent variables have long played an important role in applied research in psychology and the social sciences and, largely prompted by recent work on cognitive and non-cognitive skill formation, are becoming increasingly important in economics (Heckman, Stixrud and Urzua 2006). These models assert that a set of observable indicators is determined by a small number of common unobservable factors. The coefficients linking the latent factors and the indicators which reflect them are known as factor loadings and the configuration of the loadings is an important clue to the conceptual meaning of the unobserved factors. However, there are serious problems of interpretation in models with multiple factors, since no factor can be distinguished unambiguously from the others without further information. The indeterminacy of the estimated factors is often addressed by means of rotation algorithms, which search for combinations of factors that can be given simple interpretations. Discussions of rotation methods often take as their starting point Thurstone's (1947) analysis, which laid down the following five principles cast in terms of the properties that conceptually meaningful factor loadings are expected to display (Thurstone 1947, p.335):

(i) Every indicator should have at least one zero factor loading;

(ii) The number of zero loadings for a given factor should equal the number of factors;

(iii) For any pair of factors, there should exist several indicators which have zero loadings for one but not the other;

(iv) In models with many factors, for each pair of factors, there should exist several indicators with zero loadings for both;

(v) For any pair of factors, there should be few indicators with non-zero loadings for both.

The Thurstone principles have been found helpful, but they have no clear theoretical basis. In contrast, for many applications of factor models, we have strong a priori information on the sign of some or all of the factor loadings. For example, if the indicators are test scores and the factors are cognitive abilities, we would definitely expect non-negative loadings – it is hardly plausible to suggest that, for a given level of mathematical ability, higher verbal ability makes it more difficult to do well in a maths test. Similarly, if the observed indicators are measures of disability and the latent factors are to be interpreted as underlying physical or mental impairments, we would again expect the loadings to be non-negative. In the social sciences, these potentially powerful sign restrictions are rarely imposed on factor rotation algorithms, and non-negativity does not appear as one of Thurstone's five principles. More recent reviews have also had little to say about the signs of factor loadings (see, for example, Sass and Schmitt 2010). In the natural sciences, nonnegativity is more often exploited through Positive Matrix Factorisation algorithms (see Paatero, P. and Tapper 1994, Plumbley 2003), but the factor models typically used there are restrictive in other respects and non-negativity is exploited in a way that may not work well in the applications typical of social science, where models involve random residuals and samples are rarely large enough for sampling error to be unimportant.

2 Factor rotation

Consider the following prototypical factor model:

$$\boldsymbol{y}_i = \boldsymbol{A}^0 \boldsymbol{f}_i^0 + \boldsymbol{\eta}_i \tag{1}$$

where $\boldsymbol{y}_i$ is a $J \times 1$ vector of indicators observed for the ith of n sampled individuals and $\boldsymbol{\eta}_i$ is the corresponding vector of measurement errors. The superscript ^{0} here and elsewhere indicates that $\boldsymbol{A}^0$ is the matrix of factor loadings corresponding to the $Q \times 1$ vector of 'meaningful' latent factors $\boldsymbol{f}_i^0$. Note that (1) is incomplete and requires some specification of the process generating the factors $\boldsymbol{f}_i^0$. It can also be extended in various ways by the inclusion of intercepts and covariates in (1) or complex observation mechanisms for $\boldsymbol{y}_i$. Such extensions make no essential difference for our purposes.

Models of this class are unidentified because the structure (1) is observationally equivalent to a system with any non-singular transformation of the factors:

$$\boldsymbol{y}_i = \left(\boldsymbol{A}^0\boldsymbol{H}\right)\left(\boldsymbol{H}^{-1}\boldsymbol{f}_i^0\right) + \eta_i = \boldsymbol{A}\boldsymbol{f}_i + \eta_i \tag{2}$$

where $\boldsymbol{H}$ is an arbitrary nonsingular $Q \times Q$ matrix. This lack of identification is generally resolved by means of convenient but arbitrary normalisations during estimation of the factor loadings, followed by use of a rotation algorithm to suggest simple interpretations of the factors. For the sake of simplicity, we assume use of the following normalisation:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{A}_2 \end{pmatrix} \tag{3}$$

This normalisation involves no loss of generality provided it is possible to identify a subset of Q of the J indicators for which the corresponding rows of $\boldsymbol{A}^0$ form a nonsingular basis. In practice this requires that the basis is chosen to contain a diverse group of indicator types and, as long as this is achieved, the choice of basis is unimportant (see Pudney 1982). Note that any other set of minimal normalising assumptions (such as independent factors and orthonormal $\boldsymbol{A}$) can be recast in the form (3).

The aim of factor rotation is to find a matrix $\boldsymbol{R}$ such that $\boldsymbol{AR}$ has characteristics as close as possible to those expected of the 'meaningful' loadings matrix $\boldsymbol{A}^0$. Perfectly successful rotation would give $\boldsymbol{R} = \boldsymbol{H}^{-1} = \boldsymbol{A}_1^0$. Rotation operates by solving the following problem:

$$\min_{\boldsymbol{R}} \; \Psi\left(\hat{\boldsymbol{A}}\boldsymbol{R}\right) \tag{4}$$

where $\Psi(.)$ is a criterion function that penalises hard-to-interpret structures for the factor loading matrix and $\hat{\boldsymbol{A}}$ is an estimate of the normalised loadings matrix. Many different rotation criteria have been used, mostly members of the CF family (Crawford and Ferguson 1970):

$$\Psi\left(\boldsymbol{A}\right) = (1-\kappa) \sum_{j=1}^{J} \sum_{\substack{q,r=1 \\ q \neq r}}^{Q} a_{jq}^2 a_{jr}^2 \;+\; \kappa \sum_{q=1}^{Q} \sum_{\substack{j,k=1 \\ j \neq k}}^{J} a_{jq}^2 a_{kq}^2 \tag{5}$$

where the two components of Ψ penalise row complexity and column complexity of $\boldsymbol{A}$, respectively.

For simplicity, we assume here that all loadings are known a priori to be non-negative, but note that it is possible to exclude some of the loadings from this restriction: for example, those corresponding to a particular indicator variable. Our aim is to develop a method of rotation that incorporates a priori knowledge of the nonnegativity of $\boldsymbol{A}$, which implies inequality constraints $\boldsymbol{R} \geq \boldsymbol{0}$ and $\boldsymbol{A}_2\boldsymbol{R} \geq \boldsymbol{0}$. It is possible to impose these constraints directly:

$$\min_{\boldsymbol{R} \geq \boldsymbol{0}} \; \Psi\left(\hat{\boldsymbol{A}}\boldsymbol{R}\right) \quad \text{subject to} \quad \hat{\boldsymbol{A}}_2\boldsymbol{R} \geq \boldsymbol{0} \tag{6}$$

There are two drawbacks to this approach. First, the large number of inequality constraints make this an awkward mathematical programming problem. More important is the fact that the inequality constraints do not take account of sampling variation in the elements of $\hat{\boldsymbol{A}}_2\boldsymbol{R}$, and the solution of (6) may be dominated by restrictions on the sign of a quantity that is estimated with very low precision, leading to unstable results.[1]

A better approach is to make explicit use of a hypothesis-testing procedure which automatically takes account of sampling variability. The difficulty in doing this is that non-negativity of factor loadings involves a null hypothesis comprising multiple inequality restrictions. Wald, Likelihood Ratio and Lagrange Multiplier testing procedures for such

[1] The same criticism applies to the summations defining criteria like (5): there is a strong case for weighting the elements of these sums to take account of sampling variability, but we do not pursue that here.

hypotheses are long-established but very cumbersome computationally and lack some of the desirable properties of a good test, such as similarity on the boundary of H_0 and freedom from nuisance parameters (see Kudô 1963, Perlman 1969, Gouriéroux, Holly and Monfort 1982, Kodde and Palm 1986, Wolak 1991, Chen and Szroeter 2009). The Extreme Value test proposed by White (1990) is computationally simpler but has theoretical disadvantages. An alternative test procedure has recently been put forward by Chen and Szroeter (2009), which is extremely simple to implement and has good theoretical properties, at least asymptotically.

Our proposal for a nonnegativity-constrained rotation uses a single constraint that the set of inequality restrictions on the rotated loadings should not be rejected by the Chen-Szroeter test. Rather than using this merely as a post-rotation check (which would leave the user at an impasse in the case of a rejection), we propose incorporating the condition as part of the the rotation algorithm, in the following way:

$$\min_{\boldsymbol{R}\geq\boldsymbol{0}} \Psi\left(\hat{\boldsymbol{A}}\boldsymbol{R}\right) \quad \text{subject to} \quad P\left(\hat{\boldsymbol{A}}_2\boldsymbol{R}, \hat{\boldsymbol{V}}\right) \geq \alpha \tag{7}$$

where $P\left(\hat{\boldsymbol{A}}_2\boldsymbol{R}, \hat{\boldsymbol{V}}\right)$ is the $P-$value function of the Chen-Szroeter test, $\hat{\boldsymbol{V}}$ is an asymptotic approximation to the covariance matrix of the elements of $\hat{\boldsymbol{A}}_2$ and α is a pre-selected significance level for the non-negativity test. In most applications, the optimisation problem (7) will be simpler computationally than (6) because it involves only a single smooth inequality constraint.

3 The Chen-Szroeter nonnegativity test

For any given $\boldsymbol{R}$, write the restriction $\boldsymbol{A}_2\boldsymbol{R} \geq \boldsymbol{0}$ in vector form as follows:

$$H_0(\boldsymbol{R}): \quad \boldsymbol{G}\boldsymbol{a} \geq \boldsymbol{0} \tag{8}$$

where $\boldsymbol{G} = \boldsymbol{I} \otimes \boldsymbol{R}'$, $\boldsymbol{a} = vec(\boldsymbol{A}_2)$ and $vec(.)$ is the operator that stacks the rows of a matrix into a column vector. Writing $\boldsymbol{g} = \boldsymbol{Ga}$, the inequalities in (8) can be chained into a single equality restriction:

$$H_0(\boldsymbol{R}): \quad \sum_{s=1}^{S} g_i \, \mathbb{1}\left(g_s \leq 0\right) = 0 \tag{9}$$

where $\mathbb{1}(.)$ is the indicator function, g_s is the sth element of $\boldsymbol{g}$ and $S = (J - Q)Q$ is the number of elements of $\boldsymbol{A}_2$.

The essence of the Chen-Szroeter (CS) test is to replace the discontinuous indicator function in (9) by a smoothing function which converges asymptotically to $\mathbb{1}((g_s \leq 0)$. Write $\hat{g}_s$ as the estimate of g_s based on $\hat{\boldsymbol{A}}_2$. The CS test statistic (interpretable essentially as a $P-$value for the analogue of a t-statistic) is then constructed as:

$$P = \Phi\left(\frac{\sum_{s=1}^{S} \hat{g}_s F_s}{\sqrt{\sum_{s=1}^{S}\sum_{t=1}^{S} F_s F_t c_{st}}}\right) \tag{10}$$

where $\Phi(.)$ is the $N(0,1)$ distribution function, $F(.)$ is a smooth, strictly positive, monotonically increasing function on the real line, $F_s = F\left(\hat{g}_s\sqrt{n/\ln n}\right)$ and c_{st} is the typical element of the asymptotic approximation to the variance matrix of $\hat{\boldsymbol{g}}$, constructed as $\boldsymbol{G}\hat{\boldsymbol{V}}\boldsymbol{G}'$. We follow one of Chen and Szroeter's suggestions in using the logit smoothing function $F(x) = (1 + e^x)^{-1}$. The null hypothesis (9) is rejected against the general alternative if $P < \alpha$. Note that, in common with other tests of composite hypotheses, α is the pre-selected size of the test, in other words an upper bound on the rejection probability under H_0: the actual rejection probability at any point $\boldsymbol{R}$ depends on the true value $\boldsymbol{A}_2^0\boldsymbol{R}$.

4 An example: latent disability

Our example relates to the measurement of disability in the older population. The English Longitudinal Survey of Ageing (ELSA) (Taylor et al 2003) offers a particularly rich set of information on respondents' difficulties with the activities of daily living (ADLs) (Katz et al 1963) and instrumental activities of daily living (IADLs) (Lawton and Brody 1969). The twenty-four ELSA disability indicators are listed in Table 1. They are all binary variables and we use a logit link function to relate them to the continuous latent disability measures $\boldsymbol{f}$. Work by Morciano et al (2010) indicates a 2-factor structure for these data and we use an initial normalisation suggested by their analysis, with the loadings for the indicators recording difficulty with "walking 100 yards" and "preparing a hot meal" normalised at unity for factors 1 and 2 respectively, and the cross-loadings for those two indicators set to zero. All other indicators are allowed to have non-zero loadings on both factors. This normalisation is completely general, since any other non-degenerate matrix of factor loadings can be constructed from it by means of a suitable linear transformation. In applying factor rotation, we retain the scale normalisation $a_{11} = a_{22} = 1$ by fixing the diagonal elements of $\boldsymbol{R}$ at unity, so that optimisation is carried out with respect to the two off-diagonal elements.

The model is estimated by maximum likelihood (using *MPlus* version 6.0) for the sample of 5,145 respondents aged 65 and over who give complete answers in ELSA wave 1 (2002). The estimated loadings prior to rotation are shown in the first panel of Table 1.[2] There is a single negative loading on factor 1 and six on factor 2; these are significant in the sense that the hypothesis of non-negativity for all loadings is rejected emphatically by the CS test.

Unconstrained rotation methods tend to worsen the problem of negative loadings. Table 1 shows results from three standard methods nested within the Crawford-Ferguson family: Direct quartimin ($\kappa = 0$); Varimax ($\kappa = 1/J$); and Facparsim ($\kappa = 1$). All three methods result in the same five negative loadings for factor 1, while there are respectively 11, 6 and

[2] We do not quote test results for individual loadings, because of the serious problem of multiple comparisons that afflict the element-wise interpretation of large arrays of estimated parameters.

8 negative loadings for factor 2. For all three rotation methods, the null hypothesis of non-negative rotated loadings is rejected very strongly by the CS test. These negative loadings make interpretation very difficult: what are we to make of factor 1 which is clearly linked with difficulties with physical activity but which also appears to make it *less* likely that the individual has difficulty in preparing a hot meal, managing money or taking medication?

We incorporate the non-negativity assumption by solving the optimisation problem (7) numerically, using a test size of $\alpha = 10\%$. To check that the constrained minimum was achieved, this was done independently using two different optimisation algorithms: a quasi-Newton method (implemented in the *Gauss* constrained optimisation procedure) and the simulated annealing algorithm (Goffe et al 1994, implemented in a *Gauss* routine written by E.G.Tsionas), using a penalty function approach. Both reached the same solution without difficulty. A striking feature of the results is that the same set of rotated loadings was reached, irrespective of which variant of the Crawford-Ferguson criterion was used. Thus the result is, in this example, essentially invariant to the value of the rotation parameter κ and close to the original, unrotated loadings. This finding underlines two very important points: first, that rotation methods can make it more, rather than less, difficult to find a convincing, theoretically-valid interpretation of factor loadings. Second, that prior information on the non-negativity of factor loadings, if it exists, can have much greater identifying power than the loose Thurstone principles which are reflected in various ways by standard rotation criteria.

In this example, our first estimated factor represents disability relating mainly to physical mobility, such as getting up from a bed or chair, walking around, etc. The second factor is less clearly interpretable, but appears to be associated with activities which involve a significant element of cognition and decision-making, such as shopping, taking medicines and managing money. However, the existence of many non-negligible cross-loadings suggests that reported difficulties can often result from either physical or cognition/decision problems, so that rotation methods which tend to eliminate cross-loadings may make interpretation more difficult.

Table 1 Constrained and unconstrained factor rotation: ELSA disability indicators

	Before rotation		Unconstrained rotation						Inequality constrained*	
			$\kappa = 0$		$\kappa = J^{-1}$		$\kappa = 1$			
Indicator of difficulty with...	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
walking 100 yards	1	0	1	0.010	1	0.011	1	-0.027	1	0.070
preparing a hot meal	0	1	-0.325	1	-0.418	1	-0.446	1	0.000	1
sitting for 2 hours	1.343	-0.176	1.401	-0.162	1.417	-0.161	1.422	-0.212	1.344	-0.082
getting up from a chair	2.056	-0.277	2.146	-0.256	2.172	-0.254	2.179	-0.333	2.056	-0.133
climbing several flights of stairs	1.327	0.063	1.307	0.077	1.301	0.079	1.299	0.028	1.327	0.156
climbing a flight of stairs	1.725	0.088	1.696	0.106	1.688	0.108	1.686	0.042	1.725	0.209
stooping, kneeling, or crouching	1.809	-0.131	1.851	-0.112	1.863	-0.110	1.867	-0.180	1.809	-0.004
raising arms above shoulder level	1.073	0.023	1.066	0.034	1.064	0.035	1.063	-0.006	1.073	0.098
pulling or pushing large objects	1.502	0.230	1.428	0.246	1.406	0.248	1.400	0.190	1.502	0.336
lifting weights over 10lb.	1.341	0.280	1.250	0.294	1.224	0.295	1.216	0.243	1.341	0.374
picking up a coin	0.905	0.038	0.893	0.047	0.889	0.048	0.888	0.013	0.905	0.101
dressing	1.702	-0.119	1.741	-0.102	1.752	-0.100	1.755	-0.165	1.702	0.000
walking across a room	1.820	0.164	1.767	0.183	1.752	0.185	1.747	0.115	1.820	0.292
bathing or showering	1.539	0.048	1.523	0.064	1.519	0.065	1.518	0.006	1.539	0.155
eating & cutting up food	0.649	0.259	0.565	0.266	0.540	0.267	0.533	0.242	0.649	0.305
getting in or out of bed	2.086	-0.092	2.116	-0.070	2.125	-0.068	2.127	-0.148	2.086	0.054
using the toilet	1.870	-0.095	1.901	-0.075	1.910	-0.073	1.913	-0.145	1.870	0.036
using a map to get around	0.050	0.358	-0.066	0.358	-0.100	0.358	-0.110	0.357	0.050	0.361
shopping for groceries	1.240	0.704	1.012	0.717	0.946	0.718	0.927	0.670	1.240	0.791
making phone calls	0.026	0.345	-0.086	0.346	-0.118	0.346	-0.128	0.345	0.026	0.347
taking medications	0.015	0.551	-0.165	0.552	-0.216	0.552	-0.231	0.551	0.015	0.553
work around house or garden	1.441	0.314	1.340	0.329	1.310	0.330	1.302	0.275	1.441	0.415
managing money & paying bills	-0.139	0.602	-0.335	0.601	-0.391	0.600	-0.408	0.606	-0.139	0.592
continence	0.530	0.033	0.519	0.038	0.516	0.039	0.515	0.018	0.530	0.070
P–value (Chen-Szroeter statistic)	0.125×10^{-4}		0.736×10^{-3}		0.418×10^{-3}		0.404×10^{-5}		0.104	

Note: figures with no decimal places are constrained a priori * Non-negativity test size, $\alpha = 10\%$; rotation solutions for $\kappa = 0$, J^{-1} and 1 were identical to the accuracy shown.

5 Conclusion

In this paper, we have argued that non-negativity constraints on factor loadings are strongly supported by theoretical arguments in many or most applications of latent factor models, and that they are potentially much more powerful as sources of identifying information than the loose Thurstone (1947) principles which are often appealed to as support for conventional factor rotation methods. The difficulty with non-negativity constraints is that they are hard to incorporate in rotation algorithms and, if imposed directly without reference to sampling variability, could introduce unwelcome instability. We have instead proposed that non-negativity is introduced in the form of a requirement that, after rotation, the hypothesis of non-negativity should not be rejected by a suitable statistical test. The difficulty with this approach is that it involves testing a null hypothesis comprising multiple inequality restrictions, and standard testing LR, LM and Wald methods resulting in chi-bar-squared test statistics are too cumbersome to be used in this context. Instead, we use a much simpler test with very good asymptotic properties developed by Chen and Szroeter (2009) to incorporate non-negativity via a single smooth $P-$value constraint on a standard rotation algorithm.

Our example of survey data on self-reported disability among older people is typical of applications where non-negativity is a theoretically-supported assumption: unless this type of data were believed to be seriously misleading, it would be unreasonable to expect that increased (latent) disability of any kind would reduce reported disability. In this example, three standard rotation methods perversely increase the number of negative loadings, with as many as a third of the post-rotation loadings (in the case of quartimin rotation) being negative. In every case, the composite hypothesis of non-negativity for all loadings is rejected emphatically by the CS test. When the test is introduced as an explicit constraint on the rotation, the same result is produced irrespective of the rotation method, indicating the power of the non-negativity information. After non-negativity-constrained rotation, factor 1 represents mainly disability relating to physical mobility, while the second is associated with activities which involve an element of cognition and decision-making, such as shopping, taking medicines and managing money. The existence of many cross-loadings suggests that reported difficulty with daily activities can often result from either physical or decision-making problems.

References

[1] Chen, L.-Y. and Szroeter, J. (2009). Hypothesis testing of multiple inequalities: the method of constraint chaining. Centre for Microdata Methods and Practice: CeMMaP Working Paper CWP13/09.

[2] Crawford, C. B. and Ferguson, G. A. (1970). A general rotation criterion and its use in orthogonal rotation, *Psychometrika* **35**, 321-332.

[3] Goffe, W. L., Ferrier, G. D. and Rogers, J. (1994). Global optimization of statistical functions with simulated annealing, *Journal of Econometrics* **60**, 65-99.

[4] Gouriéroux, C., Holly, A. and Monfort, A. (1982). Likelihood ratio test, Wald test and Kuhn-Tucker test in linear mmodels with inequality constraints on the regression parameters, *Econometrica* **50**, 63-80.

[5] Heckman, J. J., Stixrud, J., Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior, *Journal of Labor Economics* **24**, 411-482.

[6] Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., Jaffe, M. W. and Cleveland, M. A. (1963). Studies of Illness in the Aged. The Index of ADL: A Standardized Measure of Biological and Psychosocial Function, *Journal of the American Medical Association* **185**, 914-919.

[7] Lawton, M. P. and Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living, *The Gerontologist* **9**, 179-186.

[8] Kodde, D. A. and Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions, *Econometrica* **54**, 1243-1248.

[9] Kudô, A. (1963). A multivariate analogue of the one-sided test, *Biometrika* **50**, 403-418.

[10] Morciano, M., Zantomio, F., Hancock, R. M. and Pudney, S. E. (2010). Disability status and older people's receipt of disability benefit in British population surveys: A multi-survey latent variable structural equation approach, University of Essex: ISER Working Paper (forthcoming).

[11] Paatero, P. and Tapper, U. (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics* **5**, 111-126.

[12] Perlman, M. D. (1969). One-sided problems in multivariate analysis, *Annals of Mathematical Statistics* **40**, 549-567.

[13] Plumbley, M. D. (2003). Algorithms for nonnegative independent component analysis, *IEEE Transactions on Neural Networks* **4**, 534-543.

[14] Pudney, S.E. (1982), Estimating latent variable systems when specification is uncertain: generalized component analysis and the eliminant method, *Journal of the American Statistical Association* **77**, 883-889.

[15] Sass, D. A. and Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis, *Multivariate Behavioural Research* **45**, 73-103.

[16] Taylor, R., Conway, L., Calderwood, L., and Lessof, C. (2003). Methodology. In: M. Marmot, J. Banks, R. Blundell, C. Lessof and J. Nazroo, (Eds), *Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing* (pp. 357-374). London: The Institute for Fiscal Studies.

[17] Thurstone L. L. (1947). *Multiple Factor Analysis*. Chicago: University of Chicago Press.

[18] White, H. (2000). A reality check for data snooping, *Econometrica* **68**, 1097-1127.

[19] Wolak, F. A. (1991). The local nature of hypothesis tests involving inequality constraints in nonlinear models, *Econometrica* **59**, 981-995.